AD_____

Award Number: DAMD17-98-1-8331

TITLE: Ultrasonic Morphological Analyzers for Breast Cancer Diagnosis

PRINCIPAL INVESTIGATOR: Frederick Lizzi, Eng. Sc.D.

CONTRACTING ORGANIZATION: Riverside Research Institute
New York, New York 10036-6991

REPORT DATE: July 2000

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2000	Final (1 Jul 98 - 30 Jun 00)

4. TITLE AND SUBTITLE
Ultrasonic Morphological Analyzers for Breast Cancer Diagnosis

5. FUNDING NUMBERS
DAMD17-98-1-8331

6. AUTHOR(S)
Frederick Lizzi, Eng. Sc.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Riverside Research Institute
New York, New York 10036-6991

E-MAIL:
lizzi@rrinyc.org

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

The goal of this research was to improve ultrasonic classification of breast lesions and guide decisions regarding biopsy requirements, especially for small lesions and those in young, dense breast, which are particularly difficult to evaluate with mammography. The research developed a set of complementary ultrasonic morphological analysis procedures (UMAPs) that analyze digital ultrasonic echo data. Each UMAP extracts a particular quantitative lesion feature that is now subjectively described; these include "echogenicity," heterogeneity," "shadowing," and lesion boundary characteristics. The set of complementary UMAP features were analyzed to derive reliable and objective lesion classification.

UMAP processing was applied to digitized radio-frequency echo data previously acquired in 119 clinical breast examinations with linear-array ultrasound systems. Anonymous ancillary patients data (including reports from subsequent biopsies) and system calibration data were stored in archival data files. Also stored were the clinicians' levels-of-suspicion that a lesion is cancerous; these LOS values were based on conventional, subjective scoring of ultrasonograms.

After refining each UMAP procedure, lesion classification was evaluated using multi-parameter discrimination procedures. ROC curves for UMAP classification had an area of 0.87 ± 0.04, which was comparable to the area (0.92 ± 0.03) of ROC curves based on the clinicians' LOS.

14. SUBJECT TERMS
Breast Cancer, Ultrasound, Ultrasonic Spectrum Analysis, Breast Lesion Classification

15. NUMBER OF PAGES
28

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

✓ Where copyrighted material is quoted, permission has been obtained to use such material.

___ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

___ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

N/A In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

X For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

N/A In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

N/A In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

N/A In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Frederic L Lizzi 28 July 2000

PI - Signature Date

Table of Contents

5. INTRODUCTION

The goal of this research is to improve the ultrasonic diagnosis of breast lesions so that physicians can objectively and reliably identify those lesions requiring biopsy. This capability would reduce the large numbers (estimated as 70%) of negative breast-lesion biopsies, avoiding unneeded patient trauma and reducing biopsy-related expenses. A series of ultrasonic parameters was defined to quantitatively describe the ultrasonic backscatter and attenuation of lesions and to quantify their boundary conformations. These objective features were computed for a set of clinical breast data by analyzing ultrasound echo data, acquired directly at the transducer; the data were collected and stored prior to this program. Feature sets were correlated with subsequent biopsy findings for more than 100 clinical cases. The utility of multiple features was evaluated using statistical tests, including Receiver Operator Characteristics (ROC). A set of objective features was identified that successfully differentiated malignant lesions from benign lesions with an ROC area of 0.87 ± 0.04.

6. BODY OF REPORT

a) Program Overview

The overall goal of this program is to determine whether ultrasonic methods could be developed to objectively identify breast lesions that do not require biopsies. Such capabilities would reduce the large number of unnecessary biopsies and thereby reduce biopsy costs and needless patient trauma and anxiety. An important aspect of our goal was to identify lesion features that could be measured objectively so that variable physician skill in interpreting clinical images would not affect decisions, and the same reliable results could be obtained throughout the United States and abroad.

To accomplish our goal, we defined quantitative ultrasonic parameters derived from radio-frequency (RF) echo signals digitally acquired directly at the transducer -- before conventional signal processing and attendant information loss has occurred. As detailed in an appended article [1], we employed calibrated spectrum analysis of these signals and then derived quantitative measures of features whose potential had previously been shown using subjective interpretation of conventional ultrasonograms. [2]

We termed our methods Ultrasonic Morphological Analysis Procedures (UMAPs). Several UMAPs were designed; each measured a specific tissue feature. Lesion properties were measured using calibrated spectrum analysis [3] of RF echo signals; we generated data files that described local two-dimensional (2-D) values of spectral parameters derived from linear-regression analysis of spectra (in dB). [4] These included spectral intercept (extrapolation to zero frequency) and midband fit (regression value at the center frequency). As described below, lesion boundaries were outlined on B-mode images and specific UMAPs were implemented to quantify: lesion echogenicity, heterogeneity, attenuation, and surface conformation descriptors.

In summary, our approved tasks were to: 1) develop software needed to compute these UMAPs; 2) refine each UMAP on selected clinical cases; 3) develop batch-mode processing software and apply the UMAPs to stored, previously acquired clinical data; 4) correlate sets of UMAPs with lesion types (benign and malignant) as specified in recorded histological reports; 5) statistically evaluate the efficacy of lesion identification using sets of UMAPs; and 6) identify the most useful UMAPs and statistically compare results with recorded physician level-of-suspicion (LOS) based on subjective interpretation of ultrasonograms and mammograms.

b) Ultrasonic System and Calibration

The RF breast data for this program had been previously acquired using an ATL HDI Ultramark 9 system with an L10-5 linear array (Advanced Technology Laboratories, Bothell, WA). Data were collected at the Yale New Haven Medical Center, Thomas Jefferson University Hospital, and the University of Cincinnati Medical Center, as an adjunct to a larger ATL study of conventional ultrasonic imaging techniques. The use of human subjects was previously approved by the Institutional Review Board at each institution in accordance with all Federal and institutional requirements. Calibration data had also been acquired using a planar reflective target and rubber blocks with diffuse distributions of embedded 10-μm glass beads as described below. As part of the current program, we extended our analysis of calibration procedures, which are required to remove system artifacts. All computations and data analysis were performed over the 5- to 9-MHz frequency band, where adequate signal-to-noise ratios were obtained.

1) Theoretical analysis

Our theoretical analysis showed that measured power spectra S(f) can be expressed as a product G H(f) D(f) A(f) T(f) where f represents frequency. G is system gain; H(f) is the transfer function of the transducer and electronic pulser/receiver; D(f) is a diffraction term associated with ultrasonic propagation; A(f) is ultrasonic attenuation in intervening tissue; and, T(f) is the desired tissue spectrum. The spectral parameters we compute are the slope, intercept, and midband value of the linear regression fit to 10 log S(f). [3] Our goal in system calibration was to determine the system function G H(f) D(f) and to remove these system artifacts so that A(f) T(f) could be computed. These procedures are described below. After correction for these system factors, we estimated A(f) and compensated for its effects as described in Sect. 3. These procedures permitted us to evaluate spectral parameters associated with the desired tissue function T(f).

Our analysis started with well-known diffraction integrals that compute ultrasonic beam profiles from specified transducer apertures. We computed transmitted beam profiles for each of the 7 transmit focal lengths provided by the ATL system with the L10-5 array. We then invoked reciprocity relationships to compute the spectrum of received backscatter echo signals when the transducer was aimed at tissue structures whose stochastic microstructure was characterized by a Gaussian autocorrelation function (describing the spatial distribution of acoustic impedance). [4]

The general approach to the problem was the same described in a previous report [4] for focussed transducers with radial symmetry. However, the current analysis had to deal with the rectangular apertures of the ATL system and the fact that ATL transducers utilize three different

focal lengths: 1) a fixed transmit/receive focal length in elevation; 2) a selectable azimuthal transmit focal length; and 3) a dynamically-varied azimuthal receive focal length.

Our analysis led to the general multiplicative result described above. We then computed theoretical spectra for specific calibration targets which were designed to provide spectra data needed to compensate for H(f) and D(f).

First, to compensate for H(f), we analyzed spectra from a planar target that was placed 20 mm from the transducer, with its surface parallel to that of the transducer. For moderately focussed beams with radial symmetry, we had previously shown that such targets provide a direct, accurate measurement of H(f). [4] For the ATL transducer, a more complex situation arose because the transducer's three focal lengths were not equal. Therefore, we used our analysis to compute correction factors that can be applied to measured planar-target spectra in order to determine H(f).

Second, to compensate for D(f), we analyzed spectra from a target comprising a rubber block that contained a diffuse distribution of 10-μm glass microspheres. We computed the power spectrum of this target as a function of range for all of the selectable transmit focal lengths. Our goal was to guide measurements of these spectra, and we mathematically demonstrated that such measurements do indeed permit accurate estimates of D(f). Our analysis indicated that sufficient sampling of D(f) was provided if measurements were taken at 5-mm intervals.

The above analyses showed how calibration-target data can be used to determine system and transducer functions (H(f) and D(f)). We also derived results that guided the efficient application of these factors to compensate clinical data for these artifacts. Specifically, since linear regression is a linear operation, we computed the slope, intercept, and midband fit of the function 10 log H(f) + 10 log D(f) and subtracted each parameter from the slope, intercept, and midband fit, respectively, of clinical spectra (10 log S(f)). This operation is much more efficient than a point-by-point division of S(f) by H(f) D(f), and it was used in our algorithms, as described below.

2) Calibration files

We implemented the calibration procedures described above in the following manner. For each of the three systems and L10-5 arrays, we had obtained RF data from an RTV surface and from rubber blocks (at 5-mm increments using ranges from 15 to 55 mm.) These data were acquired at each transmit focal length using a default energy setting kept constant in all calibration and examination procedures. We then derived spectral compensation factors for H(f) and D(f) using procedures described above; for each transducer, we entered these compensation factors in digital

files indexed by transmit focal length.

In subsequent processing, header information in each clinical RF data file specified the transducer and transmit focal length that was employed in the examination. Our algorithms used these identifiers to select the corresponding compensation file, which was automatically applied to compensate the spectral parameters of scanned tissues.

3) Spectral processing algorithms

A sequence of operations was used to process clinical data using MATLAB implementations developed during this program. Stored RF echo data from each scan were first corrected for overall system gain and time-gain compensation (TGC) used during scanning. The required information was obtained from ATL data files stored via an interface module developed by Spectrasonics Imaging, Inc. (Wayne, Pa). A standard B-mode image was then computed to guide the selection of analysis regions, as described below.

2-D spectral images were then generated to depict spectral slope, intercept, and midband fit within the scanned tissue sections. As described in the appended article [1], gain-compensated RF data along each scan line were multiplied by a sliding Hamming window and subjected to a Fast Fourier Transform (FFT). Resulting spectral magnitudes were squared, converted to dB, and analyzed with linear regression to calculate local values of the three spectral parameters. Results were stored and then corrected for system artifacts using indexed data compensation files as described in the preceding section. We also compensated each spectral parameter for attenuation [3], A(f), in intervening tissue using an attenuation coefficient of 1 dB/MHz-cm. Files of each spectral parameter were then maintained for imaging and UMAP processing of data from selected regions.

UMAP procedures are described in the appended article [1]. Briefly, the operator used a mouse to outline B-mode areas corresponding to the lesion, and a set of regions anterior, posterior, and lateral to the lesion. These regions were used to assay the following properties of each lesion:

- Echogenicity: computed as the mean spectral intercept within the lesion boundary
- Heterogeneity: computed as the standard deviation of midband fit in the lesion
- Shadowing: an attenuation measure computed from ratios of average midband fit in anterior and posterior segments (see [1]).

The computations were performed by transferring B-mode traced boundaries to each spectral-parameter file and performing the relevant computations in each region. The rationale for each feature definition is explained in [1].

4) Lesion shape descriptors

To provide complementary lesion data, we examined a series of morphometric features to characterize the surface conformation of breast lesions. We examined several candidate features which were computed using the traces of lesion boundaries described above. We also computed the area of each lesion. Among the boundary descriptors [5] we evaluated are the following, where A denotes lesion area, P denotes lesion perimeter, and D is the maximum lesion diameter:

- aspect ratio : ratio of lesion height to lesion width
- compactness : $A / 4 \pi \sqrt{D}$
- form factor : $4 \pi A / P^2$
- convexity : convex perimeter / P
- solidity : A / convex area
- fractal dimension : rate of change of the logarithm of P as the measuring unit ("ruler size") is varied. [5]

The convex perimeter and area are computed by connecting discrete boundary points with sequential lines and excluding lesion segments that extend outside of the resultant polygon. (See ref. 5.) Smooth rounded shapes result in convexity values near a maximum value of unity since the polygon closely tracks the lesion boundary. Spiculated shapes result in lower values, because the spiculations form abrupt projections that lie outside of the polygon.

5) Results

We applied the calibrated spectral features and shape descriptors described above to 119 scanned clinical cases (24 cancers and 95 non-cancers). The computations were performed using a new batch-mode MATLAB application. We then used statistical MATLAB applications to generate scatter diagrams, shown below, using summary biopsy reports to identify each lesion as cancer ("CA") or non-cancer ("non-CA"). Lastly, we applied ROC analysis [6] to a set of UMAP features using a nearest-neighbor classifier with a statistical "jack-knife" procedure.

Representative scatter diagrams I - IV are shown as Figs. 1-4. Each uses a triangle to denote conjoint properties of malignant cases (CA) and a square to indicate properties of benign cases (non-CA).

Figure 1 plots echogenicity (as defined above) vs. aspect ratio. Echogenicity for CA cases

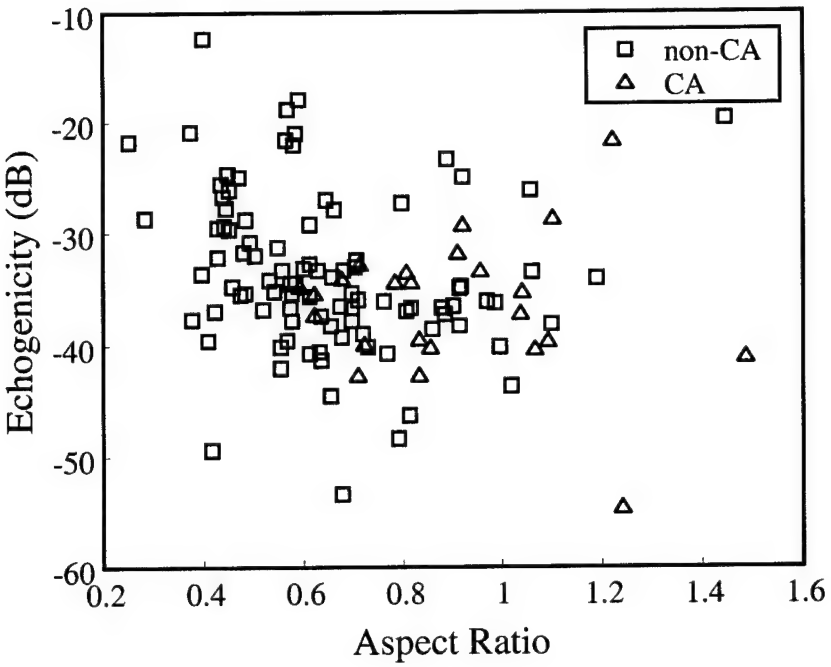


Fig. 1: Scatter diagram I.

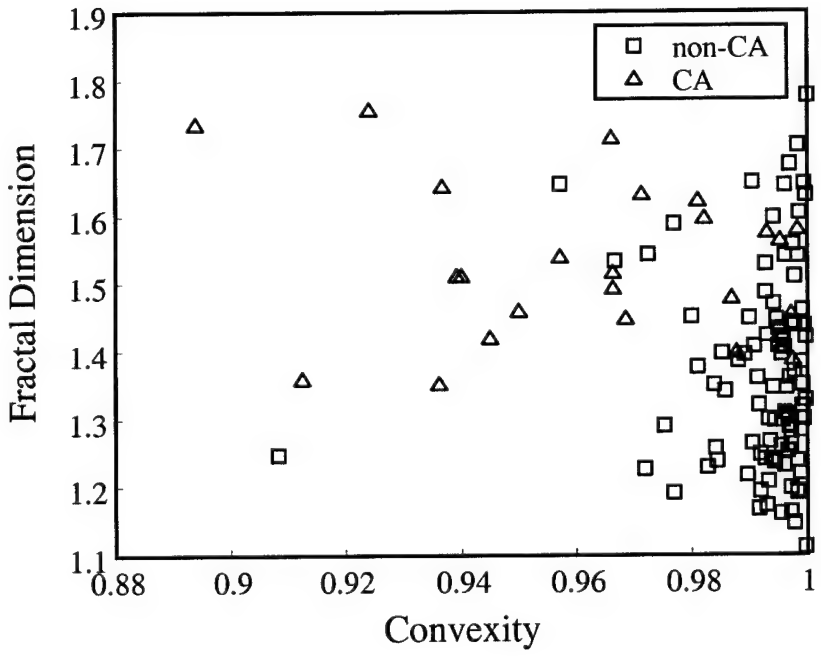


Fig. 2: Scatter diagram II.

is seen to occupy a narrow region between - 30 and - 45 dB; this range is occupied by 21 of the 24 breast cancers. The 95 non-cancers occupied a much broader range, exhibiting echogenicity levels above and below the cancer region. The mean aspect ratio for cancers (~ 1) was higher than non-cancers (~ 0.6), but large dispersions were present for both groups.

Figure 2 plots fractal dimension vs. aspect ratio. These parameters are seen to be highly correlated. Fractal dimensions, like aspect ratios, are greater in cancerous lesions.

Figure 3 plots fractal dimension vs. convexity. The mean fractal dimension of cancer cases was somewhat higher than the mean of non-cancer cases. The convexity parameter for non-cancers occupies a small regions, near the maximum possible value of unity. Cancer cases demonstrated a large spread of values, but at significantly lower values. As explained in Sect. 4, these convexity results are consistent with the smoother, rounded shapes of benign lesions as opposed to the irregular shapes of malignant lesions.

Figure 4 plots echogenicity vs. convexity. This figure clearly demonstrates the narrow central range of echogenicity for cancers and the narrow high-value range of convexity for non-cancers.

In order to quantify the diagnostic utility of UMAP features, we selected eight features for ROC analysis. [6] These features were:

1. echogenicity
2. heterogeneity
3. area
4. aspect ratio
5. form factor
6. convexity
7. solidity
8. fractal dimension

We then used a jack-knife procedure with a nearest-neighbor multiple-parameter classifier to compute an ROC curve for all 119 clinical cases. This curve, shown in Fig. 5, plots the true positive (CA) fraction (TPF) as a function of the false positive function (FPF). The figure shows that true positive factions of approximately 0.95 can be obtained at false positive fractions near 0.6. The area under the ROC curve is often used as an overall index of diagnostic utility: a random classifier exhibits an area of 0.5 while a perfect classifier exhibits an area of 1.0. As shown in the figure, the area computed for our ROC curve was 0.87 ± 0.04.

This ROC result was comparable to that derived from the level-of-suspicion (LOS) recorded by clinicians following each ultrasound examination. The clinicians had employed a numerical five-step scale ranging from high certainty that a lesion was cancerous to high certainty that it was benign.

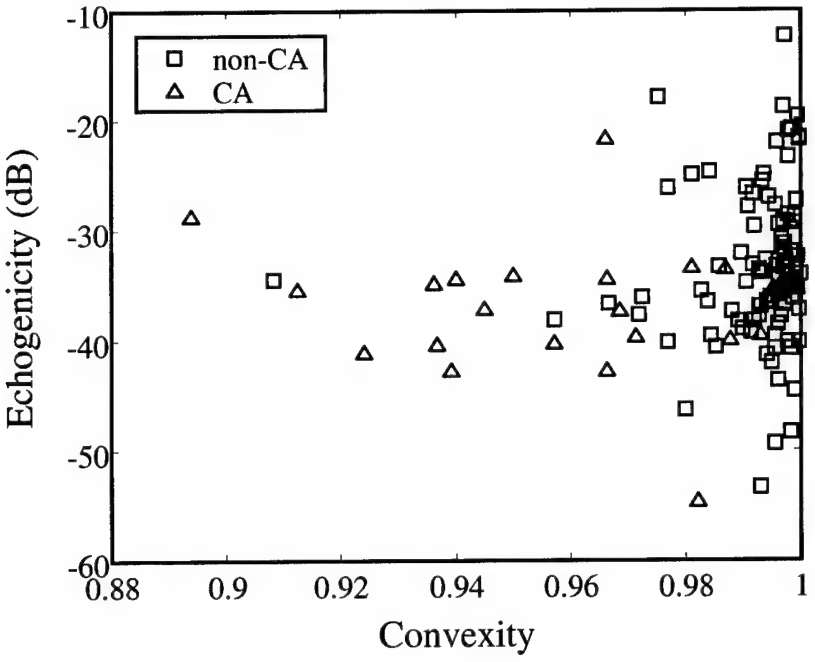


Fig. 3: Scatter diagram III.

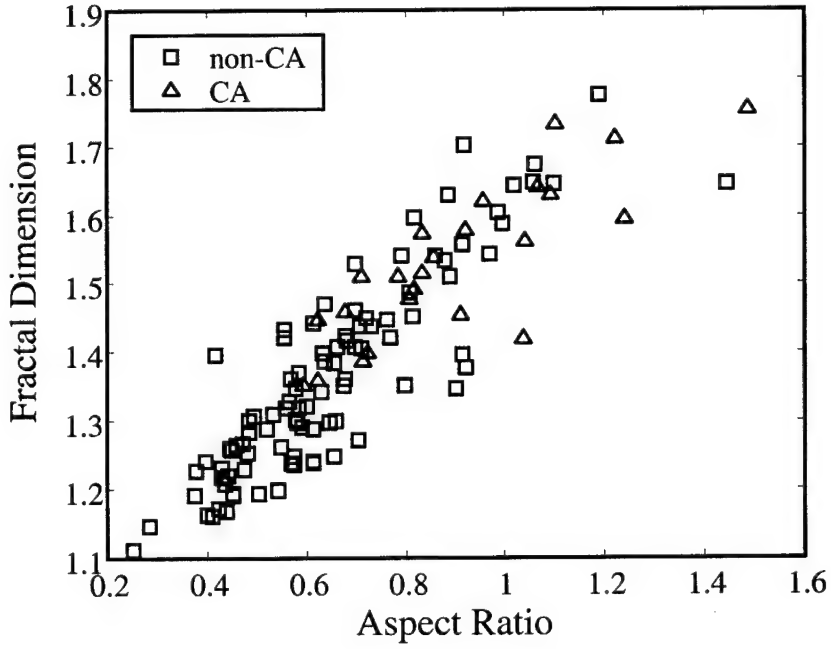


Fig. 4: Scatter diagram IV.

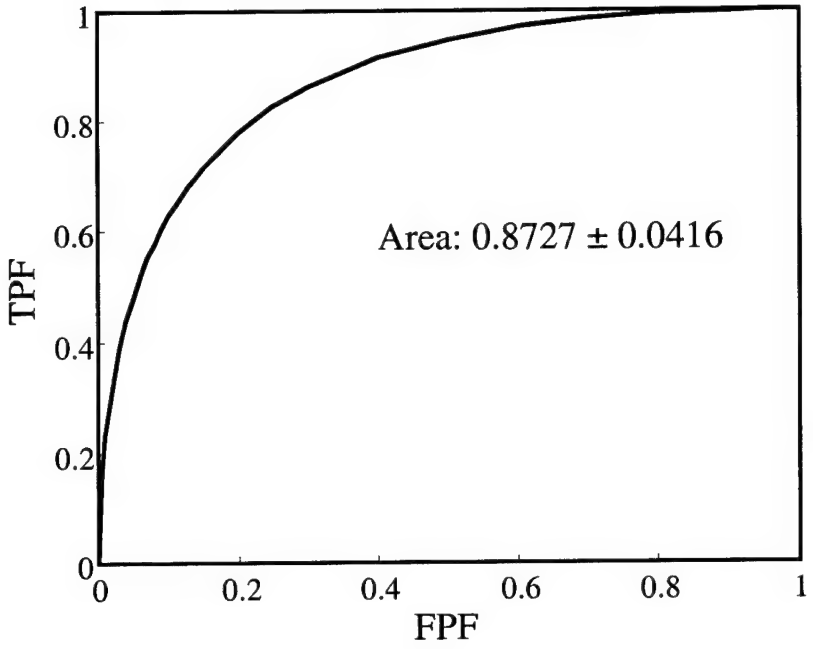


Fig. 5: ROC curve using 8 features.

In forming these conclusions, the clinicians were cognizant of additional information (mammograms, patient history, etc.). The area of the clinicians' ROC curve was 0.92 ± 0.03, close to the 0.87 ± 0.04 value we obtained using ultrasonic features. Thus, we have shown that a totally objective technique, using ultrasound data alone, can closely match the diagnostic performance of skilled clinicians who have the benefit of additional clinical information.

7. KEY RESEARCH ACCOMPLISHMENTS

- implementation of software for computing, calibrating, and displaying ultrasonic spectral features of breast tissues
- implementation of software and simple user-interface for defining breast lesions and relevant regions-of interest
- implementation and testing of software for lesion feature extraction including quantitative measures of echogenicity, heterogeneity, and attenuation
- implementation of techniques to quantify features describing the spatial configurations of lesions
- computations of above features for 24 cancerous and 95 non-cancerous breast lesions
- ROC results demonstrating an area of 0.87 ± 0.04 for differentiating cancerous and non-cancerous lesions using a specific subset of the above lesion features.

8. LIST OF REPORTABLE OUTCOMES

Conference Presentations:

"Ultrasonic Morphological Analyzers for breast cancer diagnosis," F.L. Lizzi, S.K. Alam, and E.J. Feleppa, Poster presentation at Era of Hope Conference, Atlanta, GA, June 8-11, 2000.

"Ultrasonic multifeature analysis procedures for breast lesion classification," S.K. Alam, F.L. Lizzi, E.J. Feleppa, T. Liu, and A. Kalisz, 25th International Symposium on Ultrasonic Imaging and Tissue Characterization, Arlington, VA, May 22-24, 2000.

"Ultrasonic multifeature analysis procedures for breast lesion classification," S.K. Alam, F.L. Lizzi, E.J. Feleppa, T. Liu, and A. Kalisz, American Institute of Ultrasound in Medicine (AIUM), 44th Annual Convention, San Francisco, CA, April 3-5, 2000.

"Ultrasonic multifeature analysis procedures for breast lesion classification," S.K. Alam, F.L. Lizzi, E.J. Feleppa, T. Liu, and A. Kalisz, SPIE's International Symposium on Medical Imaging: Ultrasonic Imaging and Signal Processing, San Diego, CA, February 16-17, 2000.

"Ultrasonic spectrum analysis procedures for breast cancer classification," S.K. Alam, F.L. Lizzi, E.J. Feleppa, T. Liu, and A. Kalisz, 24th International Symposium on Ultrasonic Imaging and Tissue Characterization, Arlington, VA, June 2-4, 1999.

Article

"Ultrasonic multifeature analysis procedures for breast lesion classification," S.K. Alam, F.L. Lizzi, E.J. Feleppa, T. Liu, and A. Kalisz, in *Medical Imaging 2000: Ultrasonic Imaging and Signal Processing*, K.K. Shung, M.F. Insana, (Eds.), Proceedings of SPIE, vol. 3982, pp. 196-201, 2000 (appended).

9. CONCLUSIONS

This IDEA grant showed that a set of quantitative ultrasonic descriptors has the potential for objective differentiation of breast lesions that are cancerous from those that are non-cancerous. A set of complementary features was identified that provided an ROC area (0.87) commensurate with that of skilled clinical observers, who based their decisions on ultrasound, mammography, and patient history. Further refinements of these ultrasonic parameters could thus form an objective basis for selecting lesions requiring biopsies and identifying those lesions which need not be biopsied. This would reduce unneeded biopsy procedures, avoiding unnecessary cost, trauma, and patient anxiety.

All required calibration and processing schemes were developed in this program. The full potential of the technique could be realized if future studies confirmed these results on a larger patient population, which would also provide a sufficient population size needed to better define the most efficacious set of measured features.

10. REFERENCES

[1] Alam, S.K., Lizzi, F.L., Feleppa, E.J., Liu, T., and Kalisz, A., "Ultrasonic multifeature analysis procedures for breast lesion classification," in *Medical Imaging 2000: Ultrasonic Imaging and Signal Processing*, Society of Photo-Optical Instrumentation Engineers, K.K. Shung and M.F. Insana, (Eds.), vol. 3982, pp. 196-201, Bellingham, WA, 2000.

[2] Stavros, A.E., Thickman, D., Rapp, C.L., Dennis, M.A., Parker, S.H., and Sisney, G.A., "Solid breast nodules: use of sonography to distinguish between benign and malignant lesions," Radiology, 196:123-124, 1995.

[3] Lizzi, F.L., Astor, M., Liu, T., Deng, C., Coleman, D.J., and Silverman, R.H., "Ultrasonic spectrum analysis for tissue assays and therapy evaluation," Int. J. Imag. Sys. Tech., 8:3-10, 1997.

[4] Lizzi, F.L., Greenebaum, M., Feleppa, E.J., Elbaum, M., and Coleman, D.J., "Theoretical framework for spectrum analysis in ultrasonic tissue characterization," J. Acoust. Soc. Am., 73(4):1366-1373, 1983.

[5] Russ, J.C., *Image Processing Handbook*, 3rd Ed., CRC Press, Boca Raton, FL, 1998.

[6] Metz, C.E., "ROC methodology in radiologic imaging," Inv. Radiol., 21:720-733, 1986.

11. APPENDIX

Alam, S.K., Lizzi, F.L., Feleppa, E.J., Liu, T., and Kalisz, A., "Ultrasonic multifeature analysis procedures for breast lesion classification," in *Medical Imaging 2000: Ultrasonic Imaging and Signal Processing*, Society of Photo-Optical Instrumentation Engineers, K.K. Shung and M.F. Insana, (Eds.), vol. 3982, pp. 196-201, Bellingham, WA, 2000.

In Proc. SPIE Medical Imaging 2000

Ultrasonic multifeature analysis procedures for breast lesion classification

S. K. Alam[f], F. L. Lizzi, E. J. Feleppa, T. Liu, and A. Kalisz
Riverside Research Institute, 330 West 42nd Street, New York, NY 10036

ABSTRACT

We are developing quantitative descriptors of breast lesions in order to provide reliable, operator-independent means of non-invasive breast cancer identification. These quantitative descriptors include lesion internal features assessed using spectrum analysis of ultrasonic radio-frequency (RF) echo signals and morphometric features related to lesion shape. Internal features include quantitative measures of "echogenicity," "heterogeneity," and "shadowing;" these were computed by generating spectral-parameter images of the lesion and surrounding tissue. Spectral-parameter values were generated at each pixel in the parameter image using a sliding-window Fourier analysis. Lesions were traced on B-mode images and traces were used in conjunction with spectral parameter values to compute echogenicity, heterogeneity, and shadowing. Initial results show that no single parameter may be sufficiently precise in identifying cancerous breast lesions; the results also show that the use of multiple features can substantially improve discrimination. This paper describes the background, research objective, and methodology. Clinical examples are included to illustrate the practical application of our methodology.
Keywords: Breast cancer, multifeature analysis, spectrum analysis.

1. BACKGROUND AND INTRODUCTION

Invasive ductal carcinoma comprises of 70-80% of all breast malignancies[1]. These tumors arise within milk ducts and eventually invade surrounding tissues. Of the breast biopsies (nearly 1 million annually) performed in the US, 70–90% are benign[2]. A method that reliably identifies benign lesions would reduce many unneeded biopsies, which are expensive and involve minor risks, as in any surgical procedure. Assuming the average cost of a biopsy procedure to be $2,500, even an annual reduction of 10,000 biopsies would result in a saving of a quarter billion dollars. Most breast cancers (including more than 90% of invasive ductal carcinomas[3]) are visible in conventional ultrasonic B-mode images. Recent advances in ultrasonic imaging technology present detailed examination of tumor characteristics. Although no single B-mode feature has been found to be a reliable identifier of malignancy, recent clinical studies have shown that a combination of selected features in B-mode images of breast lesions can be effective for breast cancer identification[2,4]. Diagnostically-useful characteristics included features internal to the lesions (internal features) as well as their shapes or boundaries (morphometric features). The features found to be the most useful in the multi-feature studies are listed in Table I.

Table I: Features of conventional B-mode images that are typically associated with malignant and benign lesions. A typical lesion will have only a subset of these identifying features.

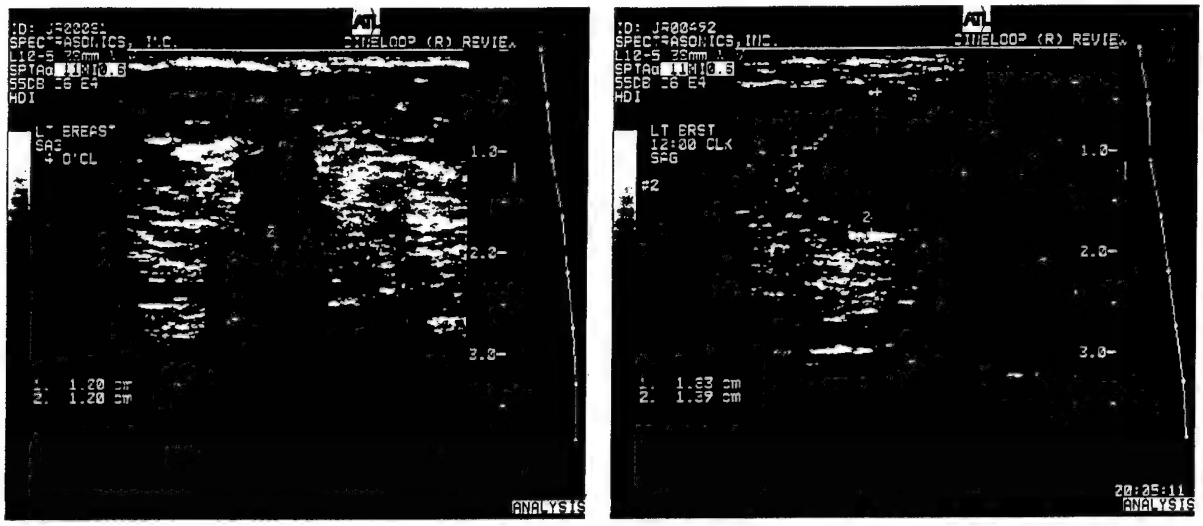
Malignant lesions		Benign lesions	
Morphometric features	Internal features	Morphometric features	Internal features
Irregular shape/spiculation	Central shadowing	Spherical/ovoid shape	Edge shadowing/enhancement
Poorly defined margin	Hypoechogenicity	Linear well-defined margin	Hyperechogenicity
Tall aspect ratio	Heterogeneous texture	Thin capsule	Homogeneous texture
Microlobulation	Calcifications	Gentle bi- or trilobulations	
Architectural distortion		Orientation parallel to tissue plane	

Figure 1 presents examples of a benign lesion and a malignant lesion with characteristics listed in Table I. Both lesions are hypoechoic. However, the malignant lesion exhibits the irregular shape, heterogeneous internal texture, and central shadow typical of malignant lesions. It also has a "tall" aspect ratio. The benign lesion exhibits a smooth shape, homogeneous internal texture, and central "anti-shadow."

Successful classification using the B-mode characteristics of Table I is invariably dependent on clinician skills. Our research addresses the development of quantitative descriptors to provide operator-independent lesion identification. Quantitative descriptors will also increase reliability of lesion identification and may allow identification of smaller lesions. As described below, we are implementing multi-feature analysis using the following internal and morphometric features.

1. Internal features: quantitative measures of echogenicity, heterogeneity, and central shadowing, based on spectrum analysis[5] of radio-frequency (RF) echoes.
2. Morphometric features: size, location, aspect ratio, and boundary smoothness.

[f] Correspondence: Email: kalam@rrinyc.org; Telephone: (212) 502-1779; Fax: (212) 502-1729



(a) (b)

Fig. 1: (a) Malignant lesion (*in situ* and invasive ductal carcinoma): The lesion has an irregular multilobular shape with a "tall" aspect ratio, heterogeneous internal texture, and a prominent posterior shadow. (b) Benign lesion (fibroadenoma): The lesion has the classical near-spherical shape, a smooth boundary, homogeneous internal texture, and a posterior "anti-shadow" or enhancement (Note the edge shadows due to refractive effects).

2. METHODS

Data Acquisition:

RF echo-signal data used in this research were digitally acquired during ultrasonic breast examinations at three separate clinical sites. At each site, a Spectrasonics Inc. (King of Prussia, PA) acquisition module was interfaced with an ATL (Bothell, WA) Ultramark 9 scanner. Data were acquired during routine ultrasonic examinations of patients scheduled for a biopsy. These patients underwent mammography prior to the ultrasound examinations and had mammographically-visible lesions. Ultrasonic data were acquired using L10-5 (7.5 MHz) linear array transducers at a default (constant) power level and a single transmit focal-length selected by the operator. Data were sampled at 20 MHz at an effective dynamic range of 14 bits. Time Gain Control (TGC) data were acquired for every scan, and the RF data were corrected for TGC before processing. All processing software was written in MATLAB™ (The Mathworks, Inc., Natick, MA), except for the tumor tracing program, which was written in Visual Basic™ (Microsoft Corporation, Redmond, WA).

Spectrum Analysis Procedure:

Calibrated spectrum analysis involves several steps. First, a Hamming window is applied to the RF data, a Fourier transform

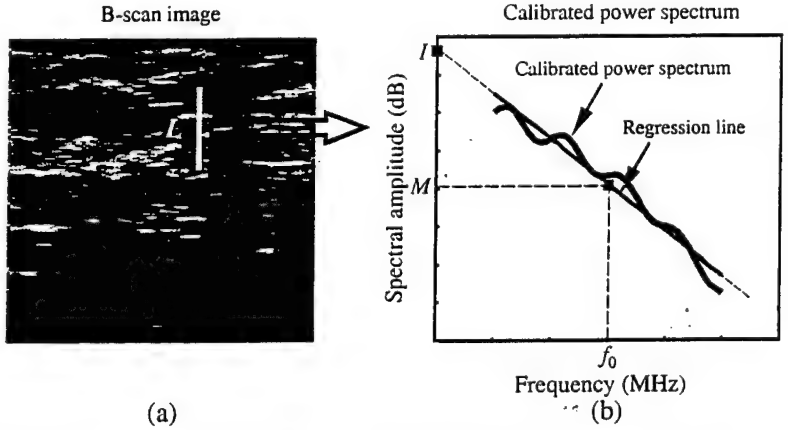
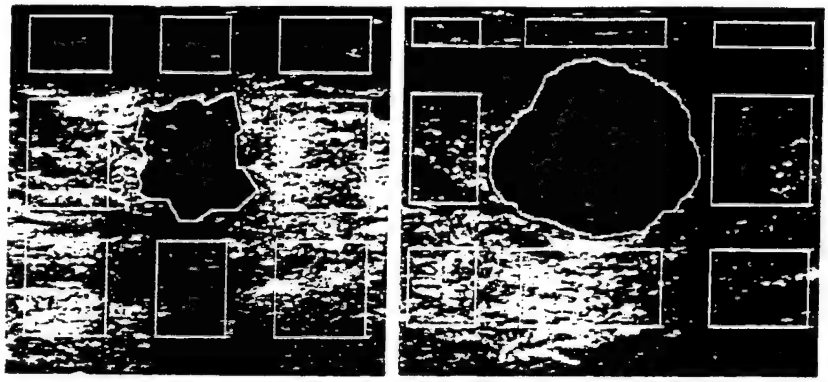


Fig. 2: Illustration of the Spectrum Analysis procedure. Calibrated power spectrum of windowed (typically Hamming, of length L) RF data is evaluated. A linear regression line through the calibrated power spectrum is computed. In this example, M is the midband value (value of the regression line at center frequency f_0) and I is the spectral intercept (value of the regression line extrapolated to $f = 0$).

is computed, and the resultant power spectrum is converted to dB. Next, system and diffraction effects[¶] are removed from the computed spectrum to derive the desired tissue spectrum. Finally, the computed spectrum is analyzed with linear regression techniques applied over the bandwidth of the signal; the primary parameters of interest are the slope of the regression line (SLP, m), its value at midpoint of signal bandwidth (MBF, M) and the intercept at zero frequency (INT, I). Images of these parameters are formed by progressively sliding the Hamming window over all RF data and repeating the above sequence. The spectrum analysis procedure is illustrated in Fig. 2.



(a) (b)

Fig. 3: B-scan images of Fig. 1 with analysis-region traces superimposed. (a) Malignant lesion. (b) Benign lesion.

Effect of Attenuation:

If no attenuation is present in intervening tissue, the linear regression line through the power spectrum is $S = I + mf$, where I is spectral intercept, m is slope, and f is frequency. The midband fit, M, is calculated as $I + mf_0$, f_0 being the center frequency. In the presence of attenuation in intervening tissue, the linear regression line through the power spectrum is $S_\alpha = S - 2\alpha Lf = I + (m - 2\alpha L)f$, where α is the effective attenuation coefficient (dB/MHz-cm) and L is the depth of intervening tissue. Thus, spectral intercept, $I_\alpha = I$, midband fit, $M_\alpha = I + (m - 2\alpha L)f_0$, and slope, $m_\alpha = (m - 2\alpha L)$. The important conclusion here is that the presence of attenuation affects slope and midband fit, but not intercept. The necessary assumption is that attenuation (in dB) varies linearly with frequency. Although this assumption is not completely accurate, the conclusion about the invariance of intercept in the presence of tissue attenuation has proved to be fairly accurate in our experience. For this report, we have corrected midband fit and slope by measuring tissue depth and using an empirical value of 1.0 dB/MHz-cm for α.

Standard Deviations of Spectral Parameters:

In homogeneous tissue regions, the standard deviations of MBF, SLP, and INT can be expressed as[7] $\sigma_M = 5.6/\sqrt{BL}$, $\sigma_s = 5.6\sqrt{12}/(B\sqrt{BL})$, and $\sigma_I = \sqrt{\sigma_M^2 + f_0\sigma_s^2}$, respectively. Here, f_0 is the center frequency (MHz), B is the bandwidth (MHz), and L is the Hamming-window length (mm). As tissue becomes more heterogeneous, the standard deviations of measured parameters increase from the above theoretical values. As noted below, we have selected σ_M to provide an index of tissue heterogeneity because it is typically much smaller than the other standard deviations; this permits smaller departures from homogeneity to be detected[†].

Internal Spectrum Analysis Features:

We have employed the following definitions to provide quantitative assays of useful but qualitative B-mode features.

1) Echogenicity is defined as the mean spectral intercept within the lesion. Because spectral intercept is independent of frequency-dependent attenuation in intervening media, no correction for attenuation is necessary.
2) Heterogeneity is defined as the standard deviation of midband fit values within the lesion. By comparing σ_M with that for a homogeneous region, the heterogeneity of a region can be assessed.
3) Shadowing is defined as the difference (normalized by lesion thickness) between mean midband fit values in comparable

[¶] Measured spectra depend not only on tissue properties, but also on 1) the combined two-way transfer function of the transducer and the ultrasonic-system electronic modules, 2) the two-way range-dependent diffraction function (beam), and 3) acoustic attenuation. Corrections for the first two functions involve experimental data obtained at each transmit focal length. The electronic transfer function was estimated from RF data acquired from the front planar surface of an RTV block in a water bath. The diffraction function was estimated from data obtained by scanning a rubber block containing a diffuse suspension of 10-μm diameter glass spheres. Finally, an empirical attenuation coefficient was used to correct for attenuation (see next subsection: *effect of attenuation*). The system effects are analyzed in detail in a separate forthcoming theoretical paper[6].

[†] For our processing parameters (L = 2.5 mm, B = 4 MHz, and f_0 = 7.5 MHz), σ_M = 1.77, σ_s = 1.53, and σ_I = 4.55.

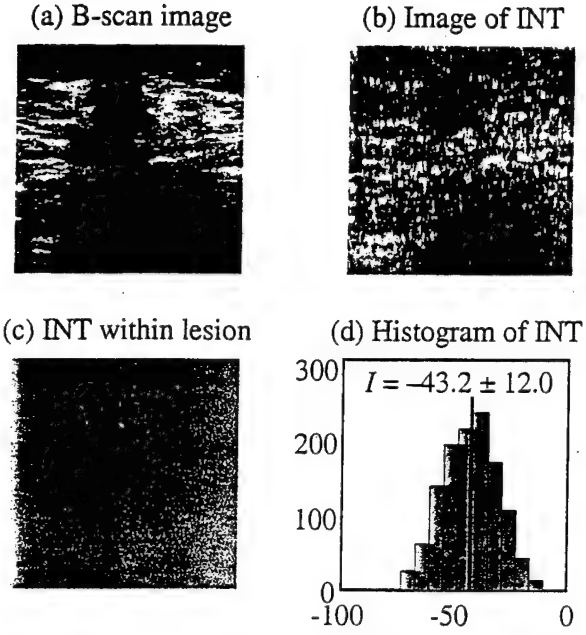


Fig. 4: Illustration of echogenicity for case of Fig. 3(a).

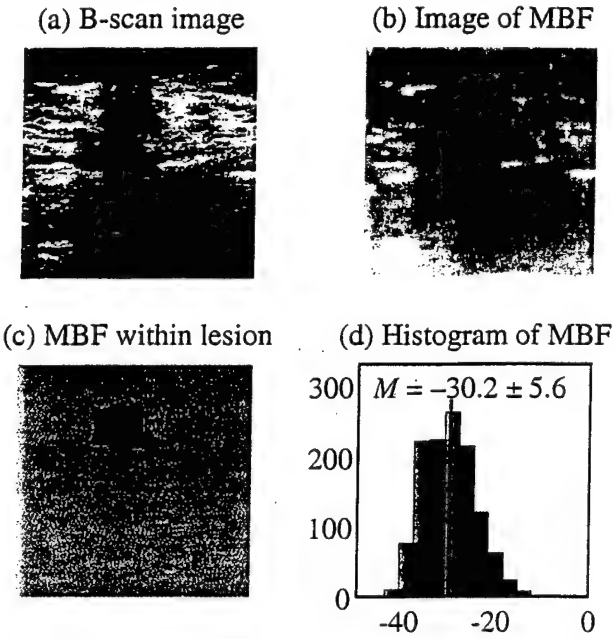


Fig. 5: Illustration of heterogeneity for case of Fig. 3(a).

shadowed and unshadowed regions posterior to the lesion. This difference can be used to estimate the attenuation coefficient within the lesion; the estimated coefficient is equal to the actual attenuation coefficient of the lesion if the shadowed and unshadowed tissues have statistically equivalent scattering properties.

We compute the three features described above by demarcating a set of analysis regions on B-mode images. Relevant spectral data from each region are then analyzed to compute the desired features. In general, nine analysis regions are used to analyze the lesion segment, anterior and posterior segments, and lateral segments, as illustrated in the following section.

Morphometric Analysis:

Invasive ductal carcinomas generally have "fuzzy" borders due to their invading margins. Cancers that have little desmoplastic reaction (proliferation of fibroblasts) typically have clear margins, but are highly irregular in shape. The aspect ratio (depth divided by width) often exceeds 0.8 in breast carcinomas for small lesions[1]. In larger carcinomas, this criterion is less useful due to their more irregular shapes and growth along duct axes. To date, among the quantitative morphometric descriptors, we have implemented lesion size (described as the total lesion area in cm^2) and aspect ratio (defined as the maximum vertical lesion-dimension divided by maximum horizontal lesion-dimension). Shape features describing irregular boundaries (or deviation from smooth shape) are crucial and are now being developed.

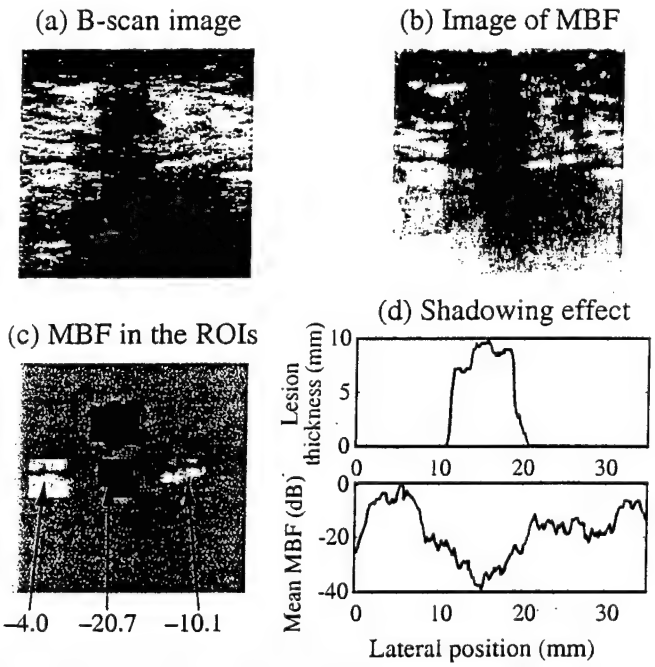


Fig. 6: Illustration of shadowing for case of Fig. 3(a).

3. RESULTS

The benign and malignant lesions of Fig. 1 are shown again in Fig. 3 with traces of the nine analysis regions superimposed. With respect to the lesion, these regions are: left-anterior, tumor-anterior, right-anterior, left-lateral, tumor, right-lateral, left-

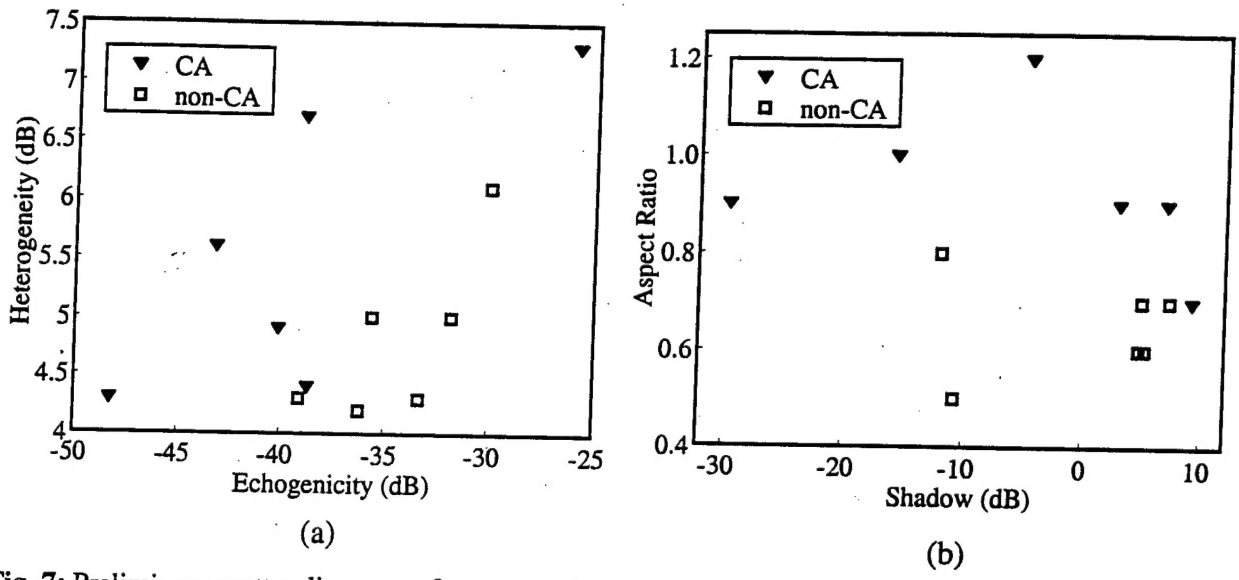


Fig. 7: Preliminary scatter diagrams of computed lesion features.

posterior, tumor-posterior, and right-posterior. For the majority of our current analysis, we need only the tumor trace; for shadow measurements, we compared mean value of MBF in the tumor-posterior region with the mean value of MBF in the left and right-posterior regions.

Estimation of lesion internal features (echogenicity, heterogeneity, and shadowing) is illustrated in Figs. 4–6 using the example of invasive ductal carcinoma *in situ* shown in Fig. 3(a). The following processing parameters were used in spectral analysis: window length, L = 2.5 mm, spectral bandwidth, B = 4 MHz (5–9 MHz), and attenuation coefficient, α = 1 dB/MHz-cm. Figure 4 illustrates the estimation of echogenicity. Figure 4(a) shows the B-scan image; the corresponding image of spectral intercept is shown in Fig. 4(b). Figure 4(c) shows the intercept image within the traced lesion boundary, and Fig. 4(d) shows the histogram of INT within the traced lesion. The quantitative value of echogenicity is the mean value of INT within the lesion (–43.2 dB). Fig. 5 shows similar images for midband fit values (MBF). The quantitative value of heterogeneity is the standard deviation of MBF inside the lesion (5.6 dB).

Figure 6 illustrates quantification of central shadowing, which is clearly visible in the B-scan image [Fig. 1(a)]. Figs. 6(a) and 6(b) are identical to Figs. 5(a) and 5(b). Fig. 6(c) shows the MBF image inside the tumor as well as in the posterior regions. Mean values of MBF in the lateral posterior regions are –4.0 and –10.1 dB (their mean being –7.05 dB). The mean of MBF posterior to the lesion is –20.7 dB, which is 13.7 dB lower. Thus, this tumor casts a 13.7 dB shadow. Fig. 6(d) separately plots the vertical thickness of the tumor and mean value of MBF in the posterior region *vs.* lateral position. The inverse relationship between lesion thickness and mean MBF is clearly visible.

In our initial studies, we have analyzed data for 12 patients (6 cancers and 6 non-cancers). Results for selected internal features and initial morphometric features are presented in the two scatter diagrams of Figs. 7(a) and (b). Although the number of patients is too small to draw any general conclusions, some promising trends are evident in both scatter diagrams. With respect to non-cancers, the cancer cases exhibit lower echogenicity, higher heterogeneity, larger aspect ratio, and larger (more negative) shadow effects.

4. CONCLUSION

We have established concept and methodology, especially for the internal features. We are in the process of completing the implementation of features describing boundary smoothness. Preliminary result for discriminating cancers from non-cancers is encouraging. This is consistent with earlier clinical results using the subjective (non-quantitative) features. Currently, the lesions are traced by non-clinicians. Prior to rigorous classification studies, lesion tracing will be performed by expert clinicians. That will be followed by multifeature statistical classification using over 100 stored patient data (after all features are implemented) to assign level of suspicion (LOS) and compare with records of clinical LOS. We anticipate automated boundary detection using sophisticated image segmentation methods in the future. A comparison with the other (contralateral) breast (possibly healthy) may yield additional insight. Because tissue properties change with time, we would include the effect of age, etc. in the analysis at a later time.

ACKNOWLEDGMENTS

This work was supported in part by US Army Medical Research and Material Command grant DAMD17-98-1-8331. Data used in this study had been acquired at the University of Cincinnati, Thomas Jefferson University, and Yale University. Ms. Rumana Huq and Ms. Stella Urban traced lesions using in-house software.

REFERENCES

1. E. Tohno, D. O. Cosgrove, and J. P. Sloane, *Ultrasound Diagnosis of Breast Diseases*, Churchill Livingstone, London, UK, 1994.
2. A. T. Stavros, D. Thickman, C. L. Rapp, M. A. Dennis, S. H. Parker, and G. A. Sisney, "Solid breast nodules: use of sonography to distinguish between benign and malignant lesions," *Radiology*, vol. 196, pp 123–134, 1995.
3. C. Cole-Beuglet, R. Z. Soriano, A. B. Kortz, B. B. Goldberg, "Ultrasound analysis of 104 primary breast carcinomas classified according to histologic type," *Radiology*, vol. 147, pp 191–196, 1983.
4. Advanced Technology Laboratories, "Summary of ATL-sponsored breast ultrasound study and preliminary findings," ATL publications, 1994.
5. F. L. Lizzi, M. Greenebaum, E. J. Feleppa, M. Elbaum, and D. J. Coleman, "Theoretical framework for spectrum analysis in ultrasonic tissue characterization," *Journal of the Acoustical Society of America*, vol. 73, pp 1366–1373, 1983.
6. F. L. Lizzi, T. Liu, S. K. Alam, and C. Deng, "Ultrasonic spectrum analysis procedures for linear arrays," *manuscript in preparation.*
7. F. L. Lizzi, E. J. Feleppa, M. Astor, and A. Kalisz, "Statistics of ultrasonic spectral parameters for prostate and liver examinations," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 44, pp 935–942, 1997.

12. BIBLIOGRAPHY

Conference Presentations:

"Ultrasonic Morphological Analyzers for breast cancer diagnosis," F.L. Lizzi, S.K. Alam, and E.J. Feleppa, Poster presentation at Era of Hope Conference, Atlanta, GA, June 8-11, 2000.

"Ultrasonic multifeature analysis procedures for breast lesion classification," S.K. Alam, F.L. Lizzi, E.J. Feleppa, T. Liu, and A. Kalisz, 25th International Symposium on Ultrasonic Imaging and Tissue Characterization, Arlington, VA, May 22-24, 2000.

"Ultrasonic multifeature analysis procedures for breast lesion classification," S.K. Alam, F.L. Lizzi, E.J. Feleppa, T. Liu, and A. Kalisz, American Institute of Ultrasound in Medicine (AIUM), 44th Annual Convention, San Francisco, CA, April 3-5, 2000.

"Ultrasonic multifeature analysis procedures for breast lesion classification," S.K. Alam, F.L. Lizzi, E.J. Feleppa, T. Liu, and A. Kalisz, SPIE's International Symposium on Medical Imaging: Ultrasonic Imaging and Signal Processing, San Diego, CA, February 16-17, 2000.

"Ultrasonic spectrum analysis procedures for breast cancer classification," S.K. Alam, F.L. Lizzi, E.J. Feleppa, T. Liu, and A. Kalisz, 24th International Symposium on Ultrasonic Imaging and Tissue Characterization, Arlington, VA, June 2-4, 1999.

Article

"Ultrasonic multifeature analysis procedures for breast lesion classification," S.K. Alam, F.L. Lizzi, E.J. Feleppa, T. Liu, and A. Kalisz, in *Medical Imaging 2000: Ultrasonic Imaging and Signal Processing*, K.K. Shung, M.F. Insana, (Eds.), Proceedings of SPIE, vol. 3982, pp. 196-201, 2000 (appended).

13. PARTICIPATING PERSONNEL

Frederic L. Lizzi, Eng.Sc.D.	Principal Investigator
Ernest J. Feleppa, Ph.D.	Co-investigator
S. Kaisar Alam, Ph.D.	"
Andrew Kalisz, M.S.	"
Tian Liu, M.S.	"
Paul Lee, M.S.	"
Angel Rosado	Technical Aide/Assistant
Rumana Huq	"